Scientific Papers. Series D. Animal Science. Vol. LXVIII, No. 2, 2025
ISSN 2285-5750; ISSN CD-ROM 2285-5769; ISSN Online 2393-2260; ISSN-L 2285-5750

STATISTICAL ANALYSIS OF COLOR FEATURES FOR QUALITY EVALUATION OF HONEY USING OPTICAL DEVICES

Eleonora NEDELCHEVA[1], Tsvetelina GEORGIEVA[1], Stanislav PENCHEV[1], Atanas ATANASOV[2], Ivaylo HRISTAKOV[2], Magdalena KACHEL[3], Plamen DASKALOV[1]

[1]University of Ruse, Automatics and Electronics, 8 Studentska Str., Ruse, Bulgaria
[2]University of Ruse, Department of Agricultural Machinery, 8 Studentska Str., Ruse, Bulgaria
[3]University of Life Sciences in Lublin, Department of Machine Operation and Production Processes Management, 13 Akademicka, Lublin, Poland

Corresponding author email: ekirilova@uni-ruse.bg

Abstract

Statistical analysis for evaluation of nine quality parameters of honey samples using two optical devices is presented in the paper. Honey from 14 regions of Bulgaria formed a set of 29 samples included in the study. The presence of heavy metals - arsenic (As), cadmium (Cd), lead (Pb), iron (Fe), Hg and pH, amount of pH, reducing sugars, sweet disaccharide and water content are preliminary evaluated in the certificated laboratory. The color digital images of the honey samples are obtained using camera of mobile phone and document camera and the data are transformed into Lab and HSV color spaces, the feature vector includes 9 color features. Correlation matrices, descriptive statistics and histograms are included in the statistical analysis. Using of document camera allow to apply more informative color features that could increase the assessment accuracy. The results show that there are informative features for quality assessment of honey samples using their color images. The identified informative signs will be integrated into developed mathematical and neural models for the quality assessment of honey through their color images.

Key words*: chemical parameters, descriptive statistics, honey, image analysis.*

INTRODUCTION

Honey production in Bulgaria is traditional as different types of organically certified honey with excellent quality indicators have been produced yearly (Dirimanova & Stoeva, 2020). In relation to distribution there are no state-owned honey production or processing entities in Bulgaria. Most of the honey is produced by small beekeepers.

Bulgaria has an international reputation as a producer of high quality honey, according to official information from Ministry of Agriculture and Food, Agrostatistics Department, Survey Beekeeping in Bulgaria (www.mzh.government.bg, 2023) the number of bee colonies on October the 1, 2023 is 816.6 thousand. The bee colonies, used for the production of honey in 2023 amounted around 659 thousand. The honey production amounts to 11 189 tons for 2023. According to the same agro-statistical survey for 2023, 4 870 tons of honey have been sold to processors, 2 392 tons have been sold directly, 483 tons – to retailers, 870 tons – to the industry, and 1 431 tons – for own consumption.

In the last few years, fake honey has started to enter the European market, and Bulgaria did not miss this trend. There are many honeys that do not meet the main quality requirements, namely: water content not more than 20 percent; electrical conductivity not more than 0.8 mS/cm; content of hydroxymethylfurfural (C6H6O3) not more than 40 mg/kg; diastase activity not less than 8 Goethe units; proline content not less than 180 mg/kg.

There are two leading methods for honey adilutration: adiluting it with glucose-fructose syrup or removing it from the hive too early. In the first case, addition of various syrups to honey increases its volume and hence reduces the production cost of the product. In the second case unripe honey has a higher water content and needs additional artificial drying. In both cases, the quality of the final product is reduced and a substance that does not meet the standards reaches the consumers.

To verify honey quality at the moment in Bulgaria, the only reliable way is testing in a certified laboratory. Every batch of honey that goes on the market goes through a series of tests in a laboratory licensed for the purpose.

Some of the mandatory tests include the following analyses:

- Presence of antibiotics, sulfonamides and depressants. Residues that appear in honey, especially as a result of the unwitting use of antibiotics to prevent bee diseases, cause food safety problems. Therefore, the European Union has a policy of zero tolerance for the presence of antibiotics in honey.
- The content of hydroxymethylfurfural (HMF) in honey is an indicator of the heating of the honey.
- Diastase activity is the main indicator proving the naturalness of honey. Prolonged exposure of bee honey to high temperatures significantly reduces the diastase activity and, accordingly, the usefulness of the honey. Mixing honey with other substances also affects this indicator, significantly reducing its value.
- Presence of impurities from commercial glucose-fructose syrup.
- Reducing sugar content- Nectar honeys must have reducing sugars higher than 68%. The average value for Bulgarian honeys is 73%.
- The pollen analysis - gives information about the percentage content of the types of pollen in the composition of a given honey and, accordingly, about its specific beneficial qualities.

There is worldwide trend to improve and facilitate food performance control methods and tools, therefore new methods are being elaborated to assist or partially replace laboratory analyses. Optical methods as tool for measurement of key indicators and quality determination provide high measurement accuracy and meet the requirements for various crop quality evaluations (Rachineni et al., 2022; Smetanska et al., 2021). Many different biological products including honey also are object of scientific interest and numerous recent surveys in the area of image analysis (Wójcik et al., 2023; Can et al., 2015; Tuberoso et al., 2014) have explored the use of different color models for product image qualification, identification of foreign objects, and more. For Bulgarian varieties of honey there are no sufficient surveys in this area.

Globally there are some works including quality assessment that deal with honey (Bogdanov et al., 2004; Dominguez & Centurión, 2015).

Al Bima et al. (Al Bima et al., 2022) have developed honey quality tester using combination of gas and color sensors. The designed tool showed good possibilities to distinguish between real honey and mixed honey for trigona honey, based on different values for RGB model.

Another survey by Wibowo et al. (Wibowo et al., 2016) is using a polarimeter and color sensor to design honey quality test equipment. The designed system is measuring color values of natural cotton honey samples.

Rochman & Mukhtar (Rochman & Mukhtar, 2019) conducted another research about the quality classification of bee honey. The proposed system is based on a single board computer and is using a spectrophotometer and machine learning approaches to classify honey.

Data analysis techniques are used to gain useful insights from datasets, which can then be used to guide future research.

Statistical analysis is the first necessary step that summarizes the data and provides the most accurate guidance for subsequent analysis.

In this study, the relationship between honey chemical content indicators and informative quantitative features derived from two optical devices is examined. Statistical analysis including descriptors from three color models (RGB, HSV and Lab) is presented in the task of honey quality control. Next these results will serve as guidance for parameter relation modelling, and creation of predicting models using colour image analysis.

MATERIALS AND METHODS

Honey from 14 regions of Bulgaria formed a set of 29 samples included in this study. The name (type) of honey samples, region (location) of production, year of production and sample code are listed as description in Table 1.

Methods for acquisition and processing include chemical analysis and image processing.

Table 1. Honey sample description, including metal content and chemical components

Name of Honey sample	Sample code	Region of production	Year of production	As	Cd	Pb	Fe	Hg	pH	Reducing sugars	Saccharose	Water
				[mg/kg]						[%]		
Rapeseed	1	Brestovica	2023	0.043	0.009	0.318	14.700	0.001	4.43±0.03	73.54±0.01	2.15±0.01	17.12±0.02
Sunflower	2	Borovo - Stara Zagora	2023	0.041	0.006	0.233	1.460	0.001	3.74±0.01	72.16±0.04	1.89±0.01	17.39±0.01
Sunflower	3	Brestovica	2023	0.041	0.006	0.264	8.690	0.001	3.79±0.01	72.36±0.03	1.83±0.03	17.38±0.02
Linden	4	Yuper	2023	0.034	0.006	0.255	6.571	0.001	4.23±0.03	73.68±0.02	2.11±0.01	17.33±0.02
Multiflower	5	Borovo-Stara Zagora	2023	0.033	0.005	0.282	7.570	0.001	3.83±0.01	74.19±0.03	2.09±0.03	17.85±0.03
Sanflower	6	Modjereto	2023	0.044	0.005	0.319	6.601	0.001	3.73±0.02	72.31±0.02	2.21±0.02	17.50±0.03
Rapeseed & Amorpha	7	Stara Zagora-Pamukchii	2024	0.049	0.006	0.304	7.190	0.001	3.66±0.03	71.15±0.04	1.31±0.02	17.13±0.02
Acacia & Mana	8	Brestowica	2023	0.044	0.006	0.339	4.191	0.001	3.92±0.02	73.15±0.01	1.31±0.01	17.09±0.03
Draka & Pustren	9	Stara Zagora	2023	0.062	0.006	0.331	4.591	0.001	3.86±0.03	75.43±0.03	1.73±0.01	17.02±0.02
Acacia	10	Yuper	2023	0.055	0.005	0.376	5.331	0.001	3.73±0.02	73.50±0.03	1.91±0.01	17.82±0.02
Multiflower	11	Yuper	2024	0.034	0.006	0.347	5.031	0.001	3.70±0.02	73.92±0.03	2.23±0.02	17.71±0.02
Multiflower	12	Yuper	2013	0.048	0.007	0.364	6.891	0.001	3.72±0.01	74.13±0.02	2.19±0.02	17.81±0.03
Sunflower	13	Trakian University	2023	0.041	0.006	0.395	6.010	0.001	3.82±0.03	72.32±0.03	1.79±0.01	17.40±0.02
Lavender	14	Stara Zagora	2022	0.042	0.006	0.398	7.003	0.001	3.48±0.03	74.47±0.04	3.25±0.026	17.52±0.03
Lavender	15	Stara Zagora	2023	0.041	0.006	0.397	5.551	0.001	3.54±0.01	74.35±0.02	3.30±0.01	17.64±0.02
Mana	16	Stara Zagora	2023	0.045	0.005	0.408	6.510	0.001	4.10±0.01	66.64±0.03	1.13±0.02	16.71±0.02
Multiflower	17	Pustrene-Stara Zagora	2023	0.048	0.008	0.385	3.550	0.001	3.75±0.01	74.43±0.03	2.07±0.05	17.64±0.01
Multiflower - nr 9	18	Razgrad	2023	0.060	0.005	0.417	5.370	0.001	3.69±0.01	74.63±0.02	2.15±0.05	17.59±0.03
Multoflower - nr 1	19	Haskovo	2023	0.043	0.005	0.448	1.541	0.001	3.67±0.02	74.52±0.03	2.11±0.01	17.45±0.01
Multiflower -nr 10	20	Ruse	2023	0.051	0.005	0.368	3.880	0.001	3.79±0.02	74.79±0.02	2.23±0.04	17.50±0.01
Multiflower - nr 8	21	Razgrad	2023	0.048	0.007	0.449	1.932	0.001	3.93±0.02	74.93±0.02	2.34±0.01	17.69±0.01
Multiflower -nr 6	22	Ruse	2023	0.064	0.006	0.408	4.650	0.001	4.14±0.02	74.89±0.01	2.49±0.02	17.63±0.04
Multiflower- nr 2	23	Turgovishte	2023	0.050	0.007	0.444	4.090	0.001	3.73±0.01	74.53±0.02	2.41±0.01	17.48±0.04
Multiflower - nr 4	24	Turgovishte	2023	0.045	0.008	0.382	0.918	0.001	3.83±0.01	74.44±0.02	2.58±0.03	17.64±0.01
Multiflower- nr 3	25	Turgovishte	2023	0.045	0.006	0.386	3.520	0.001	3.90±0.01	74.79±0.01	2.49±0.02	17.87±0.03
Multiflower - nr 6	26	Sliven	2023	0.056	0.006	0.392	4.607	0.001	3.88±0.01	74.62±0.03	2.37±0.02	17.65±0.01
Multiflower - nr 7	27	Sliven	2023	0.046	0.006	0.384	2.150	0.001	3.69±0.01	74.69±0.03	2.45±0.01	17.58±0.03
Lindon	28	Novo Village	2023	0.046	0.006	0.379	2.930	0.001	4.28±0.03	73.60±0.04	2.05±0.00	17.24±0.02
Acacia & rapeseed	29	Ivanovo	2024	0.056	0.005	0.408	2.330	0.001	3.68±0.02	72.23±0.03	1.24±0.02	16.14±0.03

Chemical analysis of honey samples was done in specified chemical laboratory conditions for determination of content of the following metals: arsenic (As), cadmium (Cd), lead (Pb), iron (Fe), mercury (Hg) and components: potential of hydrogen (pH), amount of pH, reducing sugars, sweet disaccharide and water content.

The principal methodology for chemical analysis of honey probes includes burning samples in analyser furnace, separation of mercury vapor, transformation in the form of amalgam, heating and release of free mercury, determination of components by measuring the absorption of radiation with a wavelength of 253.64 nm. AMA 254 Mercury Analyzer (Altec - Czech Republic) was used for the analysis, oxygen cylinder, nickel boats for burning samples. As software tool AMA/ama.exe analyser program was used. Our preliminary analyses (Penchev et al., 2024) include detailed presentation of laboratory results for all listed chemical parameters.

Preparation of samples for image acquisition include liquefying (decrystallization) in advance. To keep honey characteristics unaffected melting at a temperature below 45ºC was carried out. Each sample (Figure 3) contained of 10 ml decrystallized organic honey. Petri dish with 60 mm diameter was used to collect the probes.

Document Camera A405 Triumph board and smart phone Samsung Galaxy S10 Plus were used as optical devices for image acquisition of the samples. Figure 1 is showing the image acquisition setup and shooting conditions using optical devices.

The distance from the optical device to the container with the sample was 20 cm.

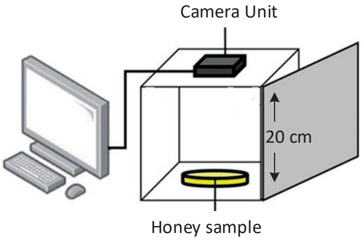


Figure 1. Image acquisition setup (original)

Next, image sample with e x d (5 x 5 cm) was cropped from each image of honey (Figure 2). All figures are original. The resulting images were stored in RGB format and used for color feature extraction.

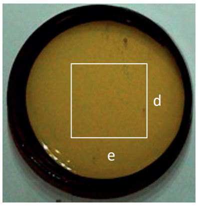
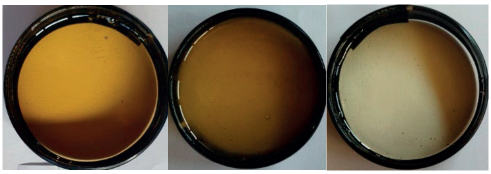


Figure 2. Cropped sample image of honey (own source)



(a)

(b)

Figure 3. Images of samples with codes 7, 17, 29 captured by: (a) phone camera and (b) document camera (own source)

The digital camera parameters and lighting typically affect digital images with RGB format that is why two more color spaces were chosen to be used further for image analysis.

Processing step includes calculation of the RGB values of each cropped image made by averaging RGB values of all pixels. Color transformations from RGB color space were performed to obtain the CIE L*a*b* and HSV color spaces.

Methods for Statistical analysis included in this study are standard and well known. Widely used statistical techniques comprise histograms, correlation matrices and also different descriptive statistics, skewness and kurtosis. For analysing data distribution characteristics skewness as fundamental approach is commonly utilized statistical analysis technique. It is used as an indicator to explain the form of data distribution, enabling an assessment of the distribution model.

The skewness was calculated by moment- based method in Excel as follows:

$$\text{Skewness} = \frac{n}{(n-1)(n-2)}\frac{\sum(x-\bar{x})^3}{s^3} \qquad (7)$$

Kurtosis provides a measurement of the extremities of the distribution of data, and therefore indicates the presence of outliers.

Excel calculates the kurtosis of a sample S as follows:

$$\text{Kurtosis} = \frac{n(n+1)\sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{\sigma}\right)^4}{(n-1)(n-2)(n-3)s^4} - \frac{3(n-1)^2}{(n-2)(n-3)} \qquad (8)$$

where $\bar{x}$ is the mean and s is the standard deviation of S. To avoid division by zero, this formula requires that $n > 3$.

When the task is to assess predictors for predicting continuous parameters, the appropriate statistical method typically depends on the complexity of the input data and the relationship between the predictors and the outcome variable.

Tests for standard normal distribution were performed in Matlab environment using one-sample Jarque-Bera test to determine if the random samples describing color components have a standard normal distribution.

Finally two-sample t-test was applied, which returns a test decision for the null hypothesis supposing data in input vectors x and y comes from independent random samples from normal distributions with equal means and equal but unknown variances. As in our case x are components derived by smart phone camera and y are components derived by document camera.

RESULTS AND DISCUSSIONS

Twenty-nine honey samples from three years of production 2022, 2023 and 2024 were analysed for As (arsenic), Cd (cadmium), Hg (mercury), Pb(lead) and Fe (iron) metal content.

Table 1 presents a report of the test results for all chemical parameters including metal content, pH, reducing sugars, sucrose and water.

The resulting values in tested samples ranged from 0.033 to 0.064 mg/kg. for As metal content. In the case of the metal Cd, content ranged from 0.005 to 0.009 mg/kg; for metal Pb ranged from 0.233 to 0.449 mg/kg, and for metal Fe ranged from 0.918 to 14.700 mg/kg. Based on the obtained results, tested honeys were characterized by low content of the metal Hg, in all tested samples it was on the border of detection.

Sugars as the main component of honey's dry matter also were analysed. For honeys in this study the average sum of reducing sugars ranged from 66.64 to 75.43%. This proves that all honeys are natural products without artificial additives (Not less than 60%).

The presence of sucrose (sweet disaccharide) and its percentage should not exceed five percent (on average, about 2-3%). It is other evidence for honey bio quality, i.e honey is not unripe, and it has not been adulterated.

All tested samples were acidic due to the obtained values, pH of honey ranged from 3.48 to 4.43; average sucrose content ranged from 1.14 to 3.30%; and the amount of water (water content) ranged from 16.14 to 17.87%.

Results in Table 1 include mean values and standard deviation for pH, reducing sugars [%], sucrose [%] and water [%].

Main descriptive statistics shown in Table 2 and Table 3 for images shoot by document camera and smart phone respectively draw the tendence for big standard deviation, big values for median, sample variance and range for most variables both for two optical devices. Only two variables S (HSV) and V (HSV) are color predictors with lowest values of standard deviation.

Table 2. Descriptive statistics for images shoot by document camera

Variable	*R*	*G*	*B*	*H*	*S*	*V*	*L*	*a*	*b*
Mean	146.75	122.25	27.98	47.69	0.82	0.58	52.15	0.47	49.09
Standard Error	2.46	3.55	4.38	0.92	0.03	0.01	1.22	1.02	1.53
Median	145.04	123.89	12.63	48.32	0.91	0.57	52.77	-0.75	52.29
Standard Deviation	18.72	27.04	33.36	6.98	0.21	0.07	9.27	7.78	11.65
Sample Variance	350.48	731.07	1112.97	48.77	0.04	0.01	85.88	60.59	135.74
Kurtosis	0.78	2.52	4.81	6.69	3.25	0.79	1.88	5.91	3.72
Skewness	-0.24	-0.80	2.00	-2.04	-1.78	-0.24	-0.59	1.90	-1.87
Range	94.21	142.93	149.78	37.60	0.87	0.37	46.84	42.73	54.56
Minimum	97.94	38.19	1.72	20.27	0.12	0.38	25.89	-12.28	9.53
Maximum	192.16	181.12	151.50	57.87	0.99	0.75	72.73	30.45	64.10

Table 3. Descriptive statistics for images shoot by smart phone

Variable	*R*	*G*	*B*	*H*	*S*	*V*	*L*	*a*	*b*
Mean	164.55	126.25	32.89	42.30	0.81	0.65	55.16	6.46	50.76
Standard Error	4.23	4.48	6.05	1.18	0.04	0.02	1.59	1.34	2.12
Median	159.59	134.08	25.25	43.30	0.86	0.63	56.60	5.96	54.40
Standard Deviation	22.80	24.10	32.58	6.37	0.19	0.09	8.57	7.21	11.43
Sample Variance	519.77	580.98	1061.73	40.61	0.04	0.01	73.45	52.01	130.69
Kurtosis	0.51	1.17	1.54	1.98	1.39	0.51	0.28	3.46	3.64
Skewness	0.41	-0.71	1.23	-1.22	-1.30	0.41	-0.38	1.27	-1.81
Range	100.63	112.92	132.19	26.64	0.74	0.39	37.65	35.57	48.35
Minimum	121.06	58.46	0.92	24.23	0.26	0.47	34.69	-4.85	14.28
Maximum	221.69	171.39	133.11	50.87	0.99	0.87	72.34	30.72	62.63

The mean values of the features for some features (R, B, a) are visibly different, while for others they have close values – S, V, b. It is necessary to test the statistical hypothesis of equality (null hypothesis) of these mean values for the samples with both optical devices. The received results using two-sample t -test at 5% significance level are described in Table 4. For the highlighted predictors in Table 4, where h=0 the samples for relevant variable have statistically equal values, p-value from t-test is also included. The result h is 1 if the test rejects the null hypothesis.

Table 4. Results from T-test at 5% significance level

Color variable	Hypothesis	Probability value
R	h=1	p=9.1871e-04
G	h=0	p=0.5265
B	h=0	p=0.6044
H	h=1	p=0.0025
S	h=0	p=0.8480
V	h=1	p=9.5055e-04
L	h=0	p=0.1758
a	h=1	p=0.0026
b	h=0	p=0.5014

Regarding computed statistics, there is positive kurtosis for all variables. The skewness and kurtosis of V (HSV) is better than all studied color predictors.

Figure 4 illustrates the skewness values for each color predictor for images shoot by smart phone and by document camera.

Approximately symmetric distributions with skewness values of -0.24 resulted for R (RGB) and V (HSV) and moderately skewed resulted -0.80 for G (RGB) were closest to zero (normal distribution). Some of components exhibited negative skewness.

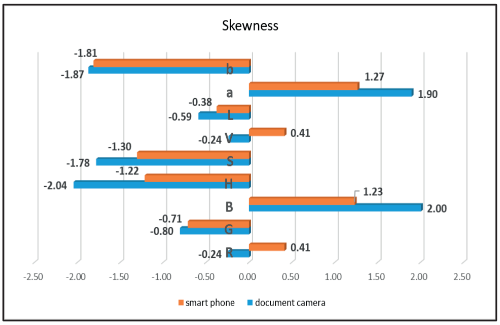


Figure 4. Skewness, images by phone and document camera

The assessment of normal distribution includes visual inspection of histograms and hypothesis testing. Figures 5 and 6 illustrate histograms of component R (RGB) for the two optical devices. Overall results showed that not all color components have normal distribution.

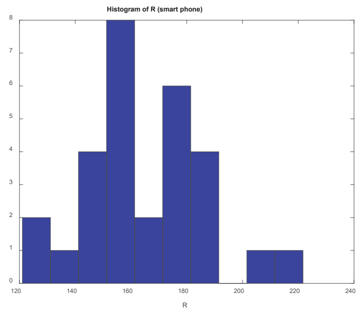


Figure 5. Histogram of component R (RGB) with smart phone camera

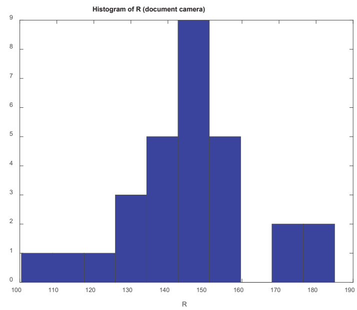


Figure 6. Histogram of component R (RGB) with document camera

Table 5 is showing test results for one-sample Jarque-Bera test at 5% significance level for all subsets.

Table 5. Test results for normal distribution of data by Jarque-Bera test at 5% significance level

Variable	Optical device	
	Document camera	Smart phone
R	h=0 p=0.5	h=0 p=0.5
G	h =1 p=0.0247	h=0 p=0.0890
B	h=1 p=1.0000e-03	h=1 p=0.0187
H	h=1 p=1.0000e-03	h=1 p=0.0154
S	h=1 p=0.0023	h=1 p=0.0170
V	h=0 p=0.5	h=0 p=0.5
L	h=0 p=0.0805	h=0 p=0.5
a	h =1 p=1.0000e-03	h=1 p=0.005
b	h=1 p=0.0010	h=1 p=0.0018

The null hypothesis says data comes from the hypothesized distribution at the significance level.

The returned value of h = 1 indicates that statistical test rejects the null hypothesis at the selected significance level. Results show that three of all variables - R, V, L have normal distribution at 5% significance level for both optical devices.

Correlation analysis was used to determine the strength of the relationship between the measured laboratory values of honey indicators and the selected color predictors. The correlations between color features and parameters were assessed using correlation matrices included in Tables 6 and 7. Marked correlations in tables are significant at $p < 0.05$.

Table 6. Correlations with color variables for images by Document camera

Variable	As	Cd	Pb	Fe	pH	Red. Sugars	Saccharose	Water
R	-0.24	-0.04	-0.40	0.37	0.19	0.29	0.21	0.24
G	-0.17	-0.04	-0.34	0.30	0.06	0.40	0.22	0.23
B	0.02	-0.31	-0.06	0.24	-0.03	-0.01	-0.18	-0.08
H	0.01	0.04	-0.18	0.03	-0.13	0.50	0.26	0.17
S	-0.05	0.33	0.02	-0.19	0.08	0.03	0.22	0.17
V	-0.23	-0.04	-0.40	0.37	0.19	0.29	0.21	0.24
L	-0.17	-0.06	-0.36	0.32	0.09	0.36	0.21	0.22
a	0.10	-0.07	0.22	-0.13	0.08	-0.49	-0.27	-0.19
b	-0.23	0.31	-0.24	0.05	0.16	0.27	0.37	0.26

Table 7. Correlations with color variables for images by smart phone

Variable	As	Cd	Pb	Fe	pH	Red. Sugars	Saccharose	Water
R	-0.23	-0.03	-0.15	0.32	-0.17	-0.01	0.08	-0.22
G	-0.05	-0.01	-0.19	0.20	-0.07	0.23	0.01	-0.24
B	-0.05	-0.23	-0.15	0.27	-0.09	-0.14	-0.37	-0.42
H	0.19	0.17	-0.04	-0.24	0.11	0.50	0.18	-0.03
S	0.00	0.27	0.15	-0.25	0.09	0.12	0.37	0.34
V	-0.23	-0.03	-0.15	0.32	-0.17	-0.01	0.08	-0.22
L	-0.10	-0.03	-0.18	0.24	-0.10	0.15	0.01	-0.26
a	-0.20	-0.11	0.14	0.12	-0.15	-0.50	-0.04	-0.03
b	-0.11	0.24	-0.04	-0.03	0.02	0.22	0.44	0.28

Correlations with color variables for images by document camera in Table 6 showed significant correlation for six of all nine examined quality parameters. There are two honey parameters - As and pH that have no significant correlations with studied color predictors. The parameter Reducing sugars has the highest value of 0.5 with H component.

Correlations with color variables for images by smart phone in Table 7 showed weak correlation at all. Despite of this H (HSV) and a (Lab) have significant correlations of 0.50 with parameter Reducing sugars in that case.

It can be concluded that better correlations between chemical parameters of honey and color predictors there are when using document camera as optical device.

In comparison with results obtained by other researchers investigating statistical analyses for data sets with biological products including honey, fruits, and juices, presented research results confirm the trend for more difficult description of the data due to the inhomogeneity of products.

In their review of statistical and machine learning methods for determination of the floral and geographical origin of honey (Maionea et al., 2019) it is concluded that multivariate data analysis and machine learning offer best performance and cost advantages for extracting valuable information from raw data sets, which can help to uncover trends and hidden patterns within data. Reviewed studies use atomic spectra and physicochemical properties, also sensorial data obtained from electronic tongue and nose, and color histograms of honey images as descriptive variables, and show high discriminative power to perform exploratory and predictive analyses. Machine learning algorithms, cluster analysis (CA) and Principal component analysis (PCA), discriminant analysis (DA) models achieve excellent results in discriminating honey origin. Study outlines the tendency of using hybrid methods that combine multivariate data analysis and machine learning techniques, as each approach has its own strengths about patterns with honey data.

Compared to the work of Tian et al. (2017) where correlations analysis of the stability, turbidity, pH and other 4 physicochemical properties for apple juices produced from nine different production technologies are determined and summarized our research show similar investigations. The results showed that one variable analysis is not enough informative for predicting the analysed properties, and the multivariate data analyses including PCA, CA and Linear Discriminant analysis have to be performed for better characterizing the qualities of juices.

The statistical approach proposed in the study of technological parameters of fruit and vegetable juices (Sotirov, 2024) showed that these analyses are extremely valuable in the area of food-processing industry. The unification of different types of products by similarities of their organoleptic indicators makes it possible to replace the different types of raw material or replace them by varieties so that they fit technologically and organoleptically, and gives good orientation on some food safety hazards.

CONCLUSIONS

In this research statistical analysis was performed using data for images of twenty-nine Bulgarian honey samples.

The relationship between digital images of honey and 8 quality indicators including presence of heavy metals - arsenic (As), cadmium (Cd), lead (Pb), iron (Fe), Hg and pH, amount of pH, reducing sugars, sweet disaccharide and water content is main subject of research. Three color models, RGB, HSV and CIE Lab, were used to describe digital images.

This study conducted tests for normal distribution, hypothesis testing, analytical information for the histograms obtained from digital images and correlation analysis.

The derived values for all statistical procedures confirm better properties of document camera as preferred optical device.

Results indicated that using document camera as optical device, variables R (RGB) and V (HSV) which have normal distributions could be used with relatively good predictive functions for predicting Pb and Fe content in honey. The reducing sugars parameter is predicted well by almost all components except S(HSV). The best predictors are H (HSV) with a correlation coefficient of 0.5 and a (Lab) with a correlation coefficient of 0.49, although they are not normally distributed. There are two parameters - As and pH that resulted in no significant correlation with no one variable.

However, based on the results obtained in this study, it is necessary to conduct additional experiments to integrate and test the identified informative signs, to develop mathematical and neural models and to analyse the practical application for quality assessment of honey through color images.

ACKNOWLEDGEMENTS

This work was supported by the Bulgarian Ministry of Education and Science under the National Research Program "Smart crop production", approved by decision of the Ministry Council, №866/26.11.2020 г.

The research was conducted within the framework of European Union-Next Generation EU, through the National Recovery and Resilience Plan of the Republic of Bulgaria, project N◦ BG-RRP-2.013–0001-C01.

REFERENCES

Al Bima, H. S., Fadli, M. N., Hidayah, N., & Arzani, L. D. P. (2022). Designing a honey quality tool based on gas sensor and color sensor. *International Journal of Engineering and Computer Science Applications, 1(*2), 94-102.

Bogdanov, S., Ruoffa, K., & Oddo, L. P. (2004). Physico-chemical methods for the characterisation of unifloral honeys: a review. *Apidologie*, (35), 4-17.

Can, Z., Yildiz, O., Sahin, H., Turumtay, E. A., Silici, S., & Kolayli, S. (2015). An investigation of Turkish honeys: Their physico-chemical properties, antioxidant capacities and phenolic profiles, *Food Chemistry*, *9*(80), 133-141. https://doi.org/10.1016/j.foodchem.2015.02.024.

Dirimanova, V., & Stoeva, T. (2020). Development of ecological beekeeping in bulgaria: status and prospects. *Scientific Papers Series Management, Economic Engineering in Agriculture and Rural Development, 20*(4), 147-151.

Dominguez, M.A., & Centurión, M. (2015). Application of digital images to determine color in honey samples from Argentina. *Microchemical Journal*, (118), 110-114. https://doi.org/10.1016/j.microc.2014.08.002.

Maionea, C., Barbosa, F., & Barbosa R.M. (2019). Predicting the botanical and geographical origin of honey with multivariate data analysis and machine learning techniques: A review, Computers and Electronics in Agriculture 157, 436-446. https://doi.org/10.1016/j.compag.2019.01.020.

Penchev, S., Georgieva, T., Manchev, G., Nedelcheva, E., Atanasov, A., Hristakov, I., Kachel, M., & Daskalov, P. (2024). Possibilities for honey quality assessment based on hyperspectral imaging (in print).

Rachineni, K., Kakita, V. M. R., Awasthi, N. P., Shirke, V. S., Hosur, V.R., & Shukla, S. C. (2022). Identifying type of sugar adulterants in honey: Combined application of NMR spectroscopy and supervised machine learning classification, *Current Research in Food Science,* (5), 272-277. https://doi.org/10.1016/j.crfs.2022.01.008.

Rochman, S., & Mukhtar, M. N. A. (2019). Classification of the quality of honey using the spectrofotometer and machine learning system based on single board computer. *Tibuana, 2*(01), 45-49. https://doi.org/10.36456/tibuana.2.01.1774.45-49

Smetanska, I., Alharthi, S. S., & Selim, K. A. (2021). Physicochemical, antioxidant capacity and color analysis of six honeys from different origin, *Journal of King Saud University - Science, 33*(5), 101447. https://doi.org/10.1016/j.jksus.2021.101447

Sotirov, A. (2024). Statistical Approach in Juice Technology, *International Journal of Scientific Development and Research, 9*(2), 761-770.

Tian, Y., Sun, L., Tanzeela, N., Liang, D., Gou, X., & Guo, Y. (2017). Multivariate statistical analysis of the quality of apple juice to integrate andsimplify juice industrial production technologies, *CYTA Journal of food*, *16*(1), 190-198. https://doi.org/10.1080/19476337.2017.1372522

Tuberoso, C.I.G., Jerković, I., Sarais, G., Congiu, F., Marijanović, Z., & Kuś, P. M. (2014). Color evaluation of seventeen European unifloral honey types by means of spectrophotometrically determined CIE L∗Cab∗hab chromaticity coordinates, *Food Chemistry,* 145, 284-291. https://doi.org/10.1016/j.foodchem.2013.08.032.

Wibowo, B. A., Rivai, M., & Tasripan, I. (2016). Alat Uji Kualitas Madu Menggunakan Polarimeter dan Sensor Warna, *Jurnal Teknik ITS, 5*(1), 28-33.

Wójcik, S., Ciepiela, F., & Jakubowska, M. (2023). Computer vision analysis of sample colors versus quadruple-disk iridium-platinum voltammetric e-tongue for recognition of natural honey adulteration, *Measurement, 209, 112514,* ISSN 0263-2241. https://doi.org/10.1016/j.measurement.2023.112514

***Ministry of Agriculture and Food, Agrostatistics Department (2024, February 22). Beekeeping in Bulgaria 2023, Retrieved from https://www.mzh.government.bg/media/filer_public/2024/06/03/ra234_beeshoney2023-eng.pdf .